ChinaXiv合作期刊
图书情报工作
第63卷 第6期 2019年3月

# 面向图书馆数字参考咨询的人机对话模型*

朱娜娜[1,2] 景东[3] 张智钧[1]

[1]哈尔滨学院图书馆 哈尔滨 150001 [2]黑龙江大学信息管理学院 哈尔滨 150001

[3]哈尔滨工业大学机电工程学院 哈尔滨 150001

**摘要：**［目的/意义］针对现有图书馆数字参考咨询人机对话机器人在对话回复内容方面的局限，提出一种融合人物画像的对话生成模型，使其回复更具个性化和趣味性，以提升图书馆智能咨询服务效果。［方法/过程］利用人机对话技术，对图书馆数字参考咨询服务中的用户和问题进行自动建模，建模方式分为个性化回复风格建模和特定用户属性建模。在个性化回复风格建模上，提出一种基于对话表示和相关性回复建模方法，该方法在学习到对话相关性的同时，利用个性化文本生成个性化的回复；在用户个人属性建模方面，基于信息抽取技术生成用户的人物画像。［结果/结论］实验结果表明，所提出的个性化回复生成模型优于已有的回复生成模型，人物画像识别的F值达到了99.8%。

**关键词：**人机对话 数字参考咨询 智能图书馆

**分类号：**G250.7

**DOI：**10.13266/j.issn.0252-3116.2019.06.001



## 1 引言

数字参考咨询服务是现代图书馆信息服务工作的一个重要组成部分。伴随着人工智能和社交媒体等信息技术的发展，数字参考咨询也迎来了新的挑战和发展机遇。

纵观人工智能行业发展，人机对话技术在研究领域和工业界受到了广泛的关注。在研究领域中，人机对话是人工智能领域皇冠上的明珠——“图灵测试”的一种实现方式，也是人工智能研究的终极目标。在工业界，Google、Facebook、苹果、微软、百度等大型跨国企业陆续推出了具备人机对话功能的应用，微软更是喊出了“对话即平台”的口号，表明了人机对话在商业上的重要性。与此同时，以API.ai、VIV.ai、WIT.ai和KITT.ai等为代表的一大批以人机对话关键技术作为核心技术的创业公司纷纷涌现，也加快了人机对话技术向产品转化的速度。因此，可以看出目前人机对话技术在研究和应用方面均具有重要的价值和意义。

在图书馆数字参考咨询服务中，能否将人机对话技术应用到现实世界中的某种载体里，如电子终端等产品，并且通过人机对话的形式，与用户在真实或虚拟世界里进行互动呢？目前，已有的数字参考咨询问答机器人一般是通过检索服务提供者预定义的FAQ数据库方式与用户进行交互，使得整个人机交互过程显得呆板和程式化，也使数字参考咨询问答机器人缺乏智能化、个性化和多样化，进而影响用户在数字参考咨询中的体验。

针对上述问题，本文利用基于深度学习的对话生成技术，从大规模背景知识中自动学习用户角色的人物属性以及个性化交互方式和内容，从而实现能够自动根据用户输入生成相关的、多样的和个性化的回复，提高问答机器人在数字参考咨询中的智能程度，使得人机对话系统不仅能够在数字参考咨询中完成咨询功能，而且还具备一定的娱乐性，从而提升用户咨询体验。

## 2 相关工作

根据术语在线（http://www.termonline.cn/index.

* 本文系国家社会科学基金项目“社交媒体突发公共事件的协同应急机制研究”（项目编号：14CXW045）和2018年黑龙江大学研究生学术交流项目“基于深度学习的开放数据与数据安全政策协同度判定”研究成果之一。

**作者简介：**朱娜娜（ORCID：0000-0002-6511-1081），馆员，博士研究生；景东（ORCID：0000-0001-9550-9595），工程师，博士研究生，通讯作者，E-mail：jingle.mail@163.com；张智钧（ORCID：0000-0003-1395-8325），馆长，教授。

收稿日期：2018-07-03 修回日期：2018-09-11 本文起止页码：5-11 本文责任编辑：王传清

htm,由全国科学技术名词审定委员会主办的术语知识服务平台)2017 年公布的定义,数字参考咨询又称虚拟参考咨询,是建立在网络基础上,不受地域、时间的限制,将用户与专家和科学专业知识联系起来的问答式咨询服务。随着网络和移动通信技术的发展和广泛应用,用户对图书馆数字参考咨询服务在响应和完成时间、答复的准确度、满意度等方面提出了更高的要求。图书馆参考咨询智能问答机器人可以实现 7 × 24 小时的实时咨询回复,能够很大程度上节省图书馆的人力、物力。但是,通过对国内 31 所省级公共图书馆的调研数据显示,目前开展网上参考咨询服务的图书馆有 26 家,能够提供网页实时资讯的有 6 家,能够提供机器人咨询的仅有浙江图书馆 1 家[1]。相比之下,高校图书馆参考咨询智能机器人的相关研究与应用较多,但主要是基于开源平台进行二次开发,如清华大学图书馆基于开源软件 A. L. I. C. E. 开发的实时智能聊天机器人“小图”[2]、上海交通大学图书馆基于 BotPlatform 开源平台构建的智能化实时咨询机器人[3]、重庆文理学院图书馆基于 AIMLBot 开源软件构建的实时虚拟参考咨询服务机器人[4]、南京大学图书馆基于微信开放的 API 接口构建的图书馆智能问答系统[5]等。现有的虚拟咨询机器人可以通过全天候、快响应的服务方式在很大程度上提高数字参考咨询服务效果,但回复内容比较刻板局限,缺乏对话的乐趣和个性化,在人机对话自动生成的智能性方面亟需提高。

人机对话系统中,机器自动回复任务旨在根据用户输入的消息,生成语法流畅语义相关的回复,因此也简称回复生成。O. Vinyals 和 A. Sordoni 等将回复生成任务抽象为一个序列到序列的学习问题[6-7]。在此基础上,L. Shang 等[8]引入了注意力机制,并进一步提出了一种混合模型。然而这类模型存在倾向于生成一般性通用回复的问题。L. Mou 等[9]将关键词作为外部知识引入生成过程以提升回复内容的多样性。C. Xing 等[10]在 L. Mou 等研究的基础上,将单关键词外部知识推广到多关键词。此外,I. V. Serban 等[11]通过在生成过程中引入一个随机的隐变量来增加回复的多样性。T. Zhao 等[12]引入条件自编码器模型,通过其中的隐变量的概率分布来建模不同回复的分布。随着回复多样性的研究推进,回复的个性化成为重要的需求。

个性化回复基于人物画像建模,其最主要的关键技术是人物属性和实体关系的抽取。早期的实体关系抽取任务来自于美国国家标准与技术研究院(National Institute of Standards and Technology,NIST)组织召开的自动内容抽取测评会议(Automatic Content Extraction,ACE)[13-16]。在此评测会议上,大多数方法是通过模式匹配或分类进行信息抽取,抽取的关系由 ACE 明确定义,属于限定域的信息抽取。随着互联网的迅速发展,限定域的信息抽取已不能满足实际的需要,因此,M. Banko[17]提出了开放域三元组抽取任务。在开放域信息抽取工作中,周蓝珺[18]分别采用基于序列模式挖掘的无指导方法和基于特征提取的有指导方法,实现了对中文音乐领域的实体关系提取。其中,基于序列挖掘的无指导方法是借助种子实体关系,在开放域检索系统中挖掘序列,进而发现新的实体关系;基于特征提取的方法则是借助字词、词性、语义角色等特征进行信息抽取。刘勇杰等[19]通过分析实体关系在句法依存树中的表述方式,提出了一种融入搜索引擎的启发式实体关系三元组抽取方法。通过使用句法依存树构造启发式模板,并在开放域检索系统中验证抽取到的实体关系是否成立。A. Fader 等[20]构建了一个信息抽取系统 ReVerb,其性能相比其他公开的信息抽取系统有显著的提升。该系统应用了词性约束与字词约束,在开放域检索系统中抽取实体及其属性,或针对特定属性关键词进行填空,用于获取实体关系。

尽管已有的对话生成模型能够生成相关性和流畅性较好的回复,人物画像的相关研究在过去也得到了较为长足的发展和进步,但是利用人物画像来建模机器人画像,从而用于生成带有特定个性化特色的人机对话系统(或聊天机器人),仍然鲜有涉猎,因此,本研究根据图书馆数字参考咨询场景中人机对话的特殊性,首次提出利用人物画像建模机器画像,构建个性化、场景化的人机对话模型。

## 3 数字参考咨询中的人机对话模型

### 3.1 对话生成模型

大规模数据驱动方法在对话系统中扮演着重要的角色,对话系统训练的大规模语料通常属于成对语料,即数据是以消息 - 回复对的形式出现的[21-22]。然而,这种成对形式的对话语料在现实生活中较难收集和扩展到较大规模。相反,非成对形式的自由文本以多种形式广泛存在,但受限于现有的训练方法,无法应用到对话系统的训练过程中。为了解决上述问题,本研究提出一种可以结合自由文本的半监督回复生成模型。一方面,大规模自由文本中存在着更为丰富的语法现象,可以提升生成回复的多样性。另一方面,还可以通过引入带有特定人物语言风格的自由文本,在回复生

成中生成具有该人物语言风格的回复。本研究将回复生成过程分解成两个子过程,利用多任务学习框架,将这两个子过程建模成两个任务分别进行学习。值得注意的是,我们将训练过程中的标准答案回复称之为模型学习的目标(简称目标),而将测试过程中模型给出的预测回复称之为回复。具体的模型结构见图1,在训练的过程(Training)中,消息(Message)和目标(Target)分别经过自编码器(AutoEncoder)进行编码学习具体的表示,这一过程称为“表示学习”过程;通过隐含语义空间(Latent Space)学习消息和目标之间的语义相关性,这一过程称为“回复相关性学习”过程。在测试阶段(Test)则直接进行消息和回复的生成测试。通过训练得到的隐含语义空间的参数保留至测试过程中。

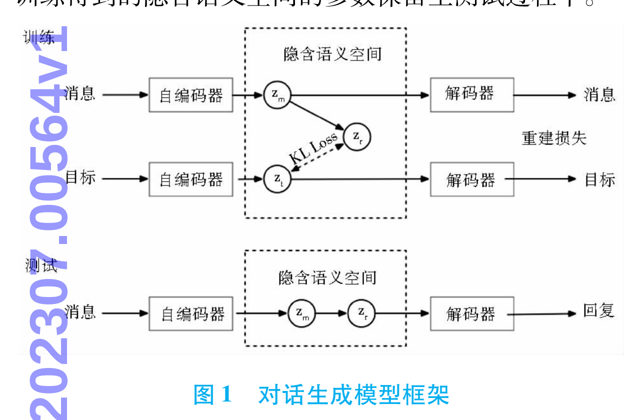



图1 对话生成模型框架

3.1.1 表示学习 对于表示学习任务,应该学习如何将一句话表示成一个模型的内部表示(向量形式),如何根据一句话的向量表示重建出这句话。这两个过程可以通过一个自编码器来完成。具体的,考虑到后续任务对于句子表示的向量空间连续性的要求,研究中使用了变分自编码器[23]。基于变分自编码器的架构,将表示句子的向量称为隐变量,其所在的空间称为隐变量空间。在这样的结构下,一旦隐变量空间学习完成,消息、模型预测的回复和标准答案的回复就可以分别被表示成空间里的3个隐变量,即消息隐变量 $z_m$、模型预测回复的隐变量 $z_r$ 和标准答案回复的隐变量 $z_t$。设计表示学习任务的损失函数如公式(1)所示:

$$L = \log P(y|x) + \mathrm{KL}(p|q) \qquad \text{公式(1)}$$

其中,$x$ 为输入,$y$ 为输出,$P$ 为序列到序列学习(Seq2seq)模型,$p$ 和 $q$ 分别表示隐变量 $z_t$ 和 $z_r$ 的概率分布,而 $z_r$ 的概率分布又是由 $z_m$ 通过隐变量空间预测出来的(见公式(2))。直观上来说,本研究希望通过计算隐变量空间中的KL距离来使得模型预测的回复与标准答案的回复尽可能接近。

3.1.2 回复相关性学习 在回复相关性学习阶段,主要是学习在隐变量空间下,如何通过消息隐变量 $z_m$ 预测一个回复隐变量 $z_r$,使得 $z_r$ 解码出的回复与标准答案在定义好的某种损失函数下的损失尽可能小。由于消息和回复已经由隐向量表示,因此,问题可以进一步转换成如何学习消息向量到回复向量的对应关系。

(1)转移网络。本研究使用一个转移网络将消息隐变量 $z_m$ 转换为回复隐变量 $z_r$,具体使用一个循环神经网络来建模该转移过程。

(2)转移关系对抗训练。由于回复生成的评价是一个开放性的问题,即使将回复与标准答案均映射到相同的隐变量空间,但是由于候选回复的多样性,仍然很难定义一个合理的损失函数用于计算消息隐变量 $z_m$ 与回复隐变量 $z_r$ 的距离。因此,本研究采用对抗训练的方式,通过引入一个判别器来计算两者之间的损失。如公式(2)和公式(3)所示:

$$z_r = f_\theta(z_m, H_m) \qquad \text{公式(2)}$$

$$\theta^* = \arg\min_\theta L_{Trans}(z_r, z_t) \qquad \text{公式(3)}$$

其中,$H_m$ 表示对消息进行自编码时的隐层状态向量,$L_{Trans}$ 是衡量模型生成的回复与标准答案回复之间向量距离的损失函数,通过对抗学习进行建模。公式(2)是通过对消息的隐层状态和隐变量建模预测出模型生成回复的隐变量。

## 3.2 人物画像建模

人物画像建模主要是在特定背景的结构化、半结构化和非结构化的文档中抽取出特定角色的属性及背景以及与其他角色之间的关系。如参考咨询员、参考咨询员与咨询者(在高校图书馆中通常为学生或教师)之间的关系以及咨询者相互之间的关系等,本文采用两种方案对人物画像进行建模,一种方案是基于模板方式;另一种方案是基于Bootstrapping的机器学习方法。

3.2.1 基于模板抽取方法

(1)Temp:直接采用人工构建的字词模板,如通过观察和总结得出“A是负责数据库资源的查询员”“B是计算机学院的学科馆员”“C咨询过B关于计算机专业书籍的馆藏信息”之类的模板。

(2)ReVerb:基于ReVerb中词性正则表达式抽取方法,ReVerb词性正则表达式如图2[20]所示:

$V \mid VP \mid VW^*P$
$V$ = verb particle? adv?
$W$ = (noun | adj | adv | pron | det)
$P$ = (prep | particle | inf. marker)

图2 Reverb中的词性正则表达式示例[20]

图书情报工作
第 63 卷 第 6 期 2019 年 3 月



（3）SRL：基于语义角色标注的抽取方法。在语义角色标注的语料上，构建如下的规则：

ATT-COO 规则：首先确定句中一个命名实体 B，找到一条指向 B 的 ATT 关系弧，对确定的结点反复寻找 ATT 关系弧确定新的结点，如果有 COO 关系弧，则进行分支。

SBV-VOB 规则：寻找 3 种结构，A、B、C、D 为 4 个按照先后顺序排列的项；A 有一条指向 B 的 SBV 关系弧，且为命名实体；C 有一条指向 D 的 ATT 关系弧；D 有一条指向 B 的 VOB 关系弧。

3.2.2 基于 Bootstrapping 的抽取方法 基于 Bootstrapping 的实体关系抽取方法，是一种利用人工构建的“种子”进行模式挖掘的方式：首先拟定某个关系或实体的一系列正确实例，即种子，然后，在历史参考咨询对话数据中检索包含种子的文本内容，将文本内容经过合适的处理，评估其可靠度，整理成一个实体或关系模板，将关系模板应用于检索系统中，再到参考咨询对话数据中进行搜索，以获得更多的种子，重复上述过程直到无法再发现新的模式。

### 3.3 融合人物画像的对话生成模型

如何在已有的对话生成模型中融入人物画像信息，从而生成包含用户属性信息的自然语言对话回复是基于人物画像的对话生成模型研究中的关键问题，本研究提出一种基于位置标记的对话生成模型，在基于神经网络的端到端的对话生成模型中，融入人物画像信息。模型框架如图 3 所示：

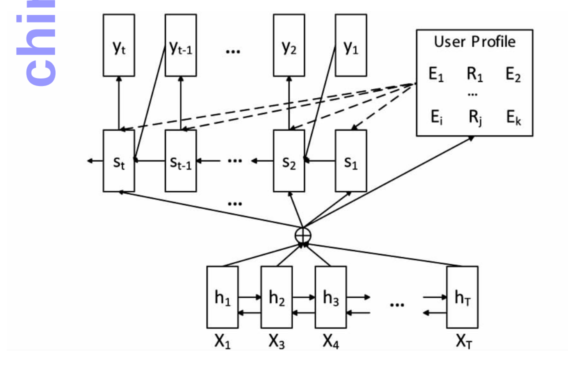



图 3 基于人物画像的对话生成模型框架

图 3 中，$X_i$（$i=1,2,\cdots,T$）表示第 i 个输入词的向量表示，$y_i$（$i=1,2,\cdots,t$）表示第 i 个预测的词，$h_i$（$i=1,2,\cdots,T$）表示第 i 个编码的隐层状态，$s_i$（$i=1,2,\cdots,t$）表示第 i 个解码的隐层状态，E 表示人物画像中的实体，R 表示人物画像中实体之间的关系，人物画像部分的三元组以隐式表示的形式参与到对话生成的解码过程中。在每一个解码过程中，对话生成模型判断当前状态下是否需要引入人物画像信息，从而决定是否引入人物画像隐式表示作为解码的输入。在输入建模的过程中采用双向 GRU 作为编码器的 RNN 模型。

## 4 实验结果及分析

为了验证所提出方法的有效性，本研究设置了对话生成、人物画像以及基于人物画像的对话生成三个方面的实验。

### 4.1 对话生成实验

本研究所提出的对话生成模型是一种领域无关的通用模型，模型结构稳定，适用于不同的数据集，并且在通用数据集（如文中的大规模微博和评论数据集）和特定领域数据集（图书馆数字参考咨询对话数据集）上，能够进行多阶段递进式训练，因此模型本身的训练是对数据不敏感的。实验采用 L. Shang 等[8]发布的大规模微博及评论数据作为对话生成的实验数据集，该数据集的具体统计情况如表 1 所示。

表 1 对话生成实验数据集统计

| | 数据类型 | 输入数（句） | 回复数（句） |
|---|---|---|---|
| 训练数据 | 1 对 1 | 1 000 000 | 1 000 000 |
| 测试数据 | 1 对多 | 1 000 | 42 422 |

本研究选取基线方法为序列到序列模型（Seq2Seq）和条件自编码器模型（CVAE）[11]。本研究的方法的第一个实验设置为使用和基线方法相同的成对语料。第二个实验设置为在成对语料的基础上引入个性化文本。4 个方法的对比实验见表 2。

表 2 对话生成实验结果

| 模型 | Avg | Ext | Gre | Distinct－1 | Distinct－2 |
|---|---|---|---|---|---|
| Seq2Seq | 0.494 | 0.354 | 0.452 | 0.002 9 | 0.017 0 |
| CVAE | 0.527 | 0.343 | 0.469 | 0.006 2 | 0.035 1 |
| 本文模型 | 0.534 | 0.376 | 0.518 | 0.011 0 | 0.073 6 |
| 本文模型＋个性化文本 | 0.543 | 0.377 | 0.513 | 0.012 7 | 0.092 6 |

在表 2 中，Avg、Ext、Gre 分别表示 Embedding 相似度计算中的平均值、单一维度极值和向量极值，这 3 个指标的详细计算方式见文献［11］。这 3 个指标越高则表示生成的回复与标准答案的语义相似度越高。Distinct－1 和 Distinct－2 为衡量生成对话中 unigram 和 bigram 多样性的指标，这两个指标越高，则表示生成的回复越具备人物的个性化说话风格。由表 2 可以看出，本研究提出的两种实验设置在 5 种指标上 Avg、Ext、Gre、Distinct－1 和 Distinct－2 均优于基线方法 Seq2Seq 和 CVAE，说明本文的方法是有效的。

### 4.2 人物画像结果

本研究采用小说数据集进行人物画像建模，该数据集的具体统计信息如表3所示：

表3 人物画像实验数据集统计

| 字符数 | 文本段数 | 主要人物数 |
|---|---|---|
| 229 315 | 2 475 | 36 |

本研究采用语言技术平台（LTP，网址为 http://ltp.ai/）对原始文本进行分词、词性标注、命名实体识别、句法分析以及语义角色标注，以提取供信息抽取所用的特征。实验结果如表4所示：

表4 人物画像关系抽取实验结果

| 方法 | Temp | ReVerb | SRL | Bootstrapping |
|---|---|---|---|---|
| 准确率 | 0.435 | 0.568 | 0.539 | 0.641 |

由表4可以看出，使用 Boostrapping 的方法超过了使用模板（Temp）、ReVerb 正则表达式和基于语义角色标注（SRL）方法的结果，其中 SRL 的结果低于 ReVerb 的结果，其主要原因是 SRL 需要命名实体识别，而命名实体识别的错误率影响了基于 SRL 方法对人物间关系的抽取结果。另外，在分析产生错误的情况时，经过统计发现，主要错误来自于"实体－关系－实体"在上下文中距离过长，导致其抽取遗漏或抽取错误。

### 4.3 基于人物画像的对话生成结果

基于人物画像的对话生成实验的参数设置见表5。在实验中，完成一轮训练大概需要10分钟，使用 Early Stopping 技术后，迭代一般在5轮左右停止。

表5 基于任务画像的对话生成实验参数设置

| 参数名称 | 参数解释 | 参数值 |
|---|---|---|
| Vocab_size | 词表大小 | 30000 |
| Max_seq_len | 序列最大长度 | 15 |
| rnn_layer | RNN 层数 | 2 |
| Embedding_dim | 词向量维度 | 200 |
| Dense_dim | MLP 隐层维度 | 256 |
| Dropout | 丢失率 | 0.3 |
| Batch_size | 批大小 | 64 |
| Epochs | 最大迭代轮数 | 10 |
| Beam_size | 束大小 | 10 |

使用验证集上准确率最高的模型作为最终的模型来预测测试集上的结果，基于人物画像的解码选择实验结果见表6。其中，评价指标采用准确率、召回率和F值。这里准确率、召回率和F值的评价对象是每个解码过程中，是否采用人物画像信息作为隐式输入，因此是一个二分类的问题。

表6 基于人物画像的解码选择实验结果

| 指标 | 分值 |
|---|---|
| 准确率 | 1.0 |
| 召回率 | 0.997 |
| F 值 | 0.998 |

由表6可知，采用对话生成模型能够非常准确地判断解码的当前状态下，是否应该引入人物画像信息。除此之外，本文继续验证嵌入人物画像信息的对话生成效果，如表7所示：

表7 基于人物画像的对话生成实验结果

| | 输入 | 标准答案回复 | 模型预测回复 | 三元组 |
|---|---|---|---|---|
| 嵌入合理 | 瑞恩·吉格斯的父亲是英格兰人么 | 呵呵，不是英国人。是威尔士的 | 是威尔士人 | （'吉格斯'，'国籍'，'威尔士'） |
| | 西游记是谁写的? | 西游记是明代吴承恩写的 | 吴承恩写的 | （'西游记'，'作者'，'吴承恩'） |
| | 2010年南非世界杯冠军应该是? | 巴西... | 当然是巴西了 | （'2010年南非世界杯'，'国籍'，'巴西'） |
| 嵌入不合理 | 江南（style）是谁唱的? | psy，鸟叔，韩国的 | 韩国唱的 | （'style'，'地区'，'韩国'） |
| | 杨幂和胡歌还会拍仙剑四吗? | 仙剑4演员换了 | 会的演员 | （'杨幂'，'职业'，'演员'） |

从表7中可以看出，人物画像三元组基本都被嵌入到了回复句中，并且大部分嵌入的是比较合理的，不过也会有一些不太适合的例子，如生成的动词不合适或者嵌入很生硬。

## 5 结语

本研究将人机对话技术应用于图书馆数字参考咨询中，通过对数字参考咨询用户的个性化回复方式和角色属性建模，实现了人机对话技术在数字参考咨询用户中的角色模拟，提升了数字参考咨询问答的智能化程度。实验结果表明本研究提出的个性化回复生成模型优于已有最好的回复生成模型。在今后的工作中，将继续研究如何将个性化回复过程与特定的人物画像信息高效结合，进一步提升对数字参考咨询用户的角色模拟能力。

**参考文献：**

[1] 郭山. 智能机器人技术在公共图书馆实时参考咨询服务中的应用[J]. 图书馆学研究,2017(10):58－61.
[2] 姚飞，纪磊，张成昱，等. 实时虚拟参考咨询服务新尝试——清华大学图书馆智能聊天机器人[J]. 现代图书情报技术，2011，27(4):77－81.

[3] 孙翌，李鲍，曲建峰. 图书馆智能化 IM 咨询机器人的设计与实现[J]. 现代图书情报技术，2011（5）：88－92.
[4] 李文江，陈诗琴. AIMLBot 智能机器人在实时虚拟参考咨询中的应用[J]. 数据分析与知识发现，2012，28（7）：127－132.
[5] 沈奎林，邵波，赵华. 利用微信构建图书馆智能问答系统[J]. 图书馆学研究，2015（8）：75－80.
[6] VINYALS O，LE Q. A neural conversational model[EB/OL]. [2015－07－22]. https://arxiv.org/pdf/1506.05869.pdf.
[7] SORDONI A，GALLEY M，AULI M，et al. A neural network approach to context-sensitive generation of conversational responses [EB/OL]. [2018－06－22]. https://arxiv.org/pdf/1506.06714.pdf.
[8] SHANG L，LU Z，LI H. Neural responding machine for short-text conversation[EB/OL]. [2018－04－27]. https://arxiv.org/pdf/1503.02364.pdf.
[9] MOU L，SONG Y，YAN R，et al. Sequence to backward and forward sequences：a content-introducing approach to generative short-text conversation[C]// Proceedings of COLING 2016，the 26th International Conference on Computational Linguistics：Technical Papers，2016：3349－3358. 2016.
[10] XING C，WU W，WU Y，et al. Topic augmented neural response generation with a joint attention mechanism[EB/OL]. [2018－09－19]. http://cn.arxiv.org/pdf/1606.08340.pdf.
[11] SERBAN I V，SORDONI A，LOWE R，et al. A hierarchical latent variable encoder-decoder model for generating dialogues[C]// Proceedings of the 39th international conference on Association for the Advancement of Artificial Intelligence. San Francisco：AAAI，2017：3295－3301.
[12] ZHAO T，ZHAO R，ESKENAZI M. Learning discourse-level diversity for neural dialog models using conditional variational autoencoders[EB/OL]. [2017－10－21]. https://arxiv.org/pdf/1703.10960.pdf.
[13] KAMBHATLA N. Combining lexical，syntactic，and semantic features with maximum entropy models for extracting relations[C]// Proceedings of the ACL 2004 on Interactive poster and demonstration sessions. Stroudsburg：Association for Computational Linguistics，2004：22.
[14] ZENG D，LIU K，LAI S，et al. Relation classification via convolutional deep neural network[C]// Proceedings of COLING 2014，the 25th international conference on computational linguistics：Technical Papers. Dublin：ACL，2014：2335－2344.
[15] LODHI H，SAUNDERS C，SHAWE-TAYLOR J，et al. Text classification using string kernels[J]. Journal of machine learning research，2002（2）：419－444.
[16] 刘克彬，李芳，刘磊，等. 基于核函数中文关系自动抽取系统的实现[J]. 计算机研究与发展，2007，4（8），1406－1411.
[17] BANKO M，CAFARELLA M J，SODERLAND S，et al. Open information extraction from the web[C]// Proceeding IJCAI '07 San Francisco：Morgan Kaufmann Publishers，2007：2670－2676.
[18] 周蓝珺. 音乐领域中文实体关系抽取研究[D]. 哈尔滨：哈尔滨工业大学，2009.
[19] 刘勇杰，姜天文，秦兵，等. 融入搜索引擎的启发式开放域三元组抽取[C]// 第五届全国知识图谱与语义计算大会. 成都：CCKS，2017.
[20] FADER A，SODERLAND S，ETZIONI O. Identifying relations for open information extraction[C]// Proceedings of the conference on empirical methods in natural language processing. Edinburgh：Association for Computational Linguistics，2011：1535－1545.
[21] BANCHS R E. Movie-DiC：a movie dialogue corpus for research and development[C]// Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics：Short Papers-Volume 2. Jeju Island：Association for Computational Linguistics，2012：203－207.
[22] LOWE R，POW N，SERBAN I，et al. The ubuntu dialogue corpus：a large dataset for research in unstructured multi-turn dialogue systems[EB/OL]. [2018－02－04]. https://arxiv.org/pdf/1506.08909.pdf.
[23] KINGMA D P，WELLING M. Auto-encoding variational bayes [EB/OL]. [2018－05－01]. https://arxiv.org/pdf/1312.6114.pdf.

**作者贡献说明：**
朱娜娜：论文选题、论文研究框架拟定、论文撰写；
景东：实验模型设计、实验分析；
张智钧：提出论文修改建议。



## A Human-Computer Dialogue Model for Digital Reference Consultation in Library

Zhu Nana[1,2]  Jing Dong[3]  Zhang Zhijun[1]
[1] Harbin University Library，Harbin 150001
[2] School of Information Management，Heilongjiang University，Harbin 150001
[3] School of Mechatronics Engineering，Harbin Institute of Technology，Harbin 150001

**Abstract**：[**Purpose/significance**] In view of the limits of response generation of conversational robots in the existing in library digital reference，this paper proposes a dialogue generation model which integrates the portraits of characters，making the reply more personalized and interesting，in order to improve the effect of library intelligent reference serv-

ice. [**Method/process**] We automatically model the specific roles and questions in digital reference service of library in two separate ways. First is to model the personalized responding style of specific role and second is to model the aspects of the role. In modeling personalized responding style, we propose an utterance representation and responding relevance-based approach to simultaneously learn the relevance of dialogue and utilize the personalized text to generate personalized responses. In modeling aspects of a specific role, we establish human profile by employing the information extraction techniques. [**Result/conclusion**] The experimental results show that, the personalized reply generation model proposed by us is superior to the best one, and the F score of user profiling recognition is 99.8%.

**Keywords**: human-computer dialogue　digital reference service　smart library